



ELSEVIER

Available online at www.sciencedirect.comLINEAR ALGEBRA
AND ITS
APPLICATIONS

Linear Algebra and its Applications 380 (2004) 151–171

www.elsevier.com/locate/laa

Generalization of Flanders' theorem to matrix triples

J. Gelonch ^{a,*,1}, C.R. Johnson ^b^a*Departamento Matemàtica, Universidad de Lleida, ETSEA, Av. Alcalde Rovira Roure 177, Lleida 25198, Spain*^b*Department of Mathematics, The College of William and Mary, Williamsburg, VA 23187-8795, USA*

Received 3 May 2001; accepted 27 May 2003

Submitted by R.A. Brualdi

Abstract

This paper deals with several ways to generalize the Flanders' theorem to matrix triples. We consider six invertible matrices and try to write them as the possible products of three matrices. Initially, we describe a wide set of necessary conditions so that this system be solvable, showing that they are not sufficient. Next, we study the simultaneous solvability of two equations, selected appropriately among the matrix system. The rest of the paper is devoted to the study of a particular case, in which the six given matrices are simultaneously diagonalizable, with distinct nonzero eigenvalues. In this case, we obtain a necessary and sufficient condition for the solvability of the full matrix system. Moreover, an explicit solution to it is constructed. Certain technical results necessary for this work may be of independent interest.

© 2003 Elsevier Inc. All rights reserved.

AMS classification: 15A21; 15A23; 15A24; 15A27; 15A30; 05A05*Keywords:* Factorization of matrices; Flanders' theorem; Permutation matrices; Commutation partition

In [1] it was first observed which pairs of matrices $B_1 \in M_m(\mathbb{C})$ and $B_2 \in M_n(\mathbb{C})$ may be written as $B_1 = A_1 A_2$ and $B_2 = A_2 A_1$ with $A_1 \in M_{m,n}(\mathbb{C})$ and $A_2 \in M_{n,m}(\mathbb{C})$. Subsequent work [8] has refined and given a straightforward proof [6] of this fundamental result. In the event that $m = n$ and B_1 and B_2 be invertible, it is necessary and sufficient that B_1 and B_2 be similar, which, by itself, is

* Corresponding author. Fax: +34-973-70-2520.

E-mail address: jgelonch@matematica.udl.es (J. Gelonch).

¹ Research partially supported by Spanish DGI grant number BFM2001-2783.

easily proven. Recently, there has been interest in similar characterization of products among $k \geq 3$ matrices. Here, because the number of possible products, using each of the k matrices exactly once, grows factorially in k , several analogous questions are possible. For example, in [2,3], the k -tuples B_1, \dots, B_k (not necessarily invertible) which may be written as cyclic products among A_1, \dots, A_k , not necessarily square, (e.g., $B_1 = A_1 A_2 \cdots A_k$, $B_2 = A_2 \cdots A_k A_1$, \dots , $B_k = A_k A_1 \cdots A_{k-1}$) were characterized. Again, the square, invertible case is relatively straightforward. Here, we consider two other natural generalizations and concentrate upon the square invertible case (which is much more complicated in these events) for $k = 3$.

Let $B_1, B_2, \dots, B_6 \in M_n(\mathbb{C})$ be given invertible matrices and consider the (non-linear) matrix equations

$$\begin{aligned} 1. B_1 &= A_1 A_2 A_3, & 2. B_2 &= A_2 A_3 A_1, & 3. B_3 &= A_3 A_1 A_2, \\ 4. B_4 &= A_1 A_3 A_2, & 5. B_5 &= A_2 A_1 A_3, & 6. B_6 &= A_3 A_2 A_1 \end{aligned} \quad (1)$$

in unknown matrices $A_1, A_2, A_3 \in M_n(\mathbb{C})$. It is clearly necessary for a solution that A_1, A_2 and A_3 be invertible. One natural generalization of Flanders' question (which we consider) is the simultaneous solvability of all six equations. Of course, now with six data matrices and only three variable matrices, solvability should impose quite stringent conditions upon B_1, B_2, \dots, B_6 . In particular, our assumption that the B_i 's (and therefore the A_i 's) are square of the same size is necessary because of the requirements of conformability of the multiplication of the A_i 's. We may also consider subsets of two or more of the six equations of the system (1). To this end, we first observe certain necessary conditions that flow from our equations.

First of all, we have certain similarities among the B_i 's via A_j 's (which are assumed to exist):

$$\begin{aligned} A_1^{-1} B_1 A_1 &= B_2, & A_2^{-1} B_2 A_2 &= B_3, & A_3^{-1} B_3 A_3 &= B_1, \\ A_1^{-1} B_4 A_1 &= B_6, & A_2^{-1} B_5 A_2 &= B_4, & A_3^{-1} B_6 A_3 &= B_5. \end{aligned} \quad (2)$$

This implies, in particular, that

$$\begin{aligned} B_1, B_2 \text{ and } B_3 &\text{ are mutually similar,} \\ &\text{i.e., have a common Jordan canonical form,} \end{aligned} \quad (3)$$

$$\begin{aligned} B_4, B_5 \text{ and } B_6 &\text{ are mutually similar,} \\ &\text{i.e., have a common Jordan canonical form,} \end{aligned} \quad (4)$$

$$B_1 \text{ and } B_4 \text{ are simultaneously similar to } B_2 \text{ and } B_6 \text{ (via } A_1), \quad (5)$$

$$B_2 \text{ and } B_5 \text{ are simultaneously similar to } B_3 \text{ and } B_4 \text{ (via } A_2), \text{ and} \quad (6)$$

$$B_3 \text{ and } B_6 \text{ are simultaneously similar to } B_1 \text{ and } B_5 \text{ (via } A_3). \quad (7)$$

In fact, some 3-tuple of matrices conveying the simultaneous similarities would have to give a simultaneous solution to Eqs. (1) or (2).

There are a number of other simultaneous similarity conditions involving various words in the B_i 's and their inverses, but all seem to be implied by (5)–(7). In view of (3) and (4), we may write

$$B_2 = C_2^{-1} B_1 C_2, \quad B_3 = C_3^{-1} B_1 C_3, \quad B_5 = C_5^{-1} B_4 C_5, \quad B_6 = C_6^{-1} B_4 C_6. \quad (8)$$

Of course the Jordan canonical form of B_1 and the one of B_4 need not be the same (examples are easily given), but both Jordan canonical forms have the same determinant, as

$$\det B_i = \det B_j, \quad \text{all } 1 \leq i, j \leq 6. \quad (9)$$

All the necessary conditions mentioned thus far do *not* characterize the simultaneous solvability of (1), as shown by the following.

Example 1. Let

$$B_1 = \begin{bmatrix} 1/2 & 0 \\ 0 & 2 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}, \quad (10)$$

$$B_4 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad B_5 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad B_6 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \quad (11)$$

Then, $\det B_i = 1$, all i , and

- $B_1 \sim B_2$ and $B_4 \sim B_6$ via $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$,
- $B_2 \sim B_3$ and $B_5 \sim B_4$ via the identity matrix and
- $B_3 \sim B_1$ and $B_6 \sim B_5$ via $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

In fact, up to scalar multiples, a calculation shows that these are the only similarities that work. Thus, all mentioned necessary conditions are met for a solution to (1).

But the matrices $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, the identity matrix and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ are a commuting family, so that they (or any scalar multiples) cannot be a solution of (1).

Moreover, if we consider the three matrices (10), we can see that the other three matrices (11) such that the system (1) is consistent, must have the form

$$B_4 = \begin{bmatrix} a & 0 \\ 0 & 1/a \end{bmatrix}, \quad B_5 = \begin{bmatrix} a & 0 \\ 0 & 1/a \end{bmatrix}, \quad B_6 = \begin{bmatrix} 1/a & 0 \\ 0 & a \end{bmatrix}.$$

However, these necessary conditions characterize simultaneous solvability for two natural types of subsets of our equations: transpositions (e.g., equations 1 and 4 of (1)) and 3-cycles (e.g., equations 1–3 of (1)). The case of 3-cycles is essentially covered in [3], but we review it, with a complete description of all solutions, for

future use (observe that the simultaneous solution of the system (1) essentially involves coincidence of solutions for the two 3-cycles, 1, 2, 3 and 4, 5, 6).

First of all, simultaneous solution of equations 1–3 of (1) is characterized by (3) and that of equations 4–6 of (1) by (4). For sufficiency, we note that

$$A_1 = B_1 C_2, \quad A_2 = C_2^{-1} C_3 \quad \text{and} \quad A_3 = C_3^{-1} \quad (12)$$

constitute a solution to equations 1–3 of (1), while

$$A_1 = B_4 C_6, \quad A_2 = C_5^{-1} \quad \text{and} \quad A_3 = C_6^{-1} C_5 \quad (13)$$

constitute a solution to equations 4–6 of (1) (the matrices C_2 , C_3 , C_5 and C_6 satisfy (8)). We note here that the situation is completely analogous for a single longer cycle.

In fact, for the case of a 3-cycle, we may characterize *all* solutions. For an n -by- n matrix B , we denote by $\mathcal{C}(B)$ the group of invertible matrices that commute with B . Given two similar matrices A and B , denote the set of invertible matrices C such that $C^{-1}AC = B$ by $\mathcal{C}(A, B)$, the set of matrices that convey the similarity. It is then a straightforward calculation that

Lemma 2. *If $A, B \in M_n(\mathbb{C})$ are similar and $Z \in \mathcal{C}(A, B)$, then $Z' \in \mathcal{C}(A, B)$ if and only if $Z' = XZY$, in which $X \in \mathcal{C}(A)$ and $Y \in \mathcal{C}(B)$.*

Proof. First, suppose that $Z, Z' \in \mathcal{C}(A, B)$. Then, from $Z^{-1}AZ = B = (Z')^{-1}AZ'$, we have $Z'Z^{-1}A = AZ'Z^{-1}$. Thus, $Z'Z^{-1} \in \mathcal{C}(A)$. We define $X = Z'Z^{-1}$ and $Y = I$, to obtain $Z' = XZY$ with $X \in \mathcal{C}(A)$ and $Y \in \mathcal{C}(B)$.

Conversely, suppose that we can write $Z' = XZY$ with $X \in \mathcal{C}(A)$ and $Y \in \mathcal{C}(B)$. Then,

$$\begin{aligned} (Z')^{-1}AZ' &= (XZY)^{-1}A(XYZ) = Y^{-1}Z^{-1}X^{-1}AXZY \\ &= Y^{-1}Z^{-1}AZY = Y^{-1}BY = B, \end{aligned}$$

as desired. \square

We may then observe

Lemma 3

- (a) *If $A_1 = X_1$, $A_2 = X_2$ and $A_3 = X_3$ constitute a simultaneous solution to equations 1–3 of (1), then $A_1 = \tilde{X}_1$, $A_2 = \tilde{X}_2$ and $A_3 = \tilde{X}_3$ constitute another solution if and only if*

$$\tilde{X}_1 = T_1^{-1}X_1T_2, \quad \tilde{X}_2 = T_2^{-1}X_2T_3 \quad \text{and} \quad \tilde{X}_3 = T_3^{-1}X_3T_1, \quad (14)$$

in which $T_i \in \mathcal{C}(B_i)$, $i = 1, 2, 3$.

- (b) *If $A_1 = Y_1$, $A_2 = Y_2$ and $A_3 = Y_3$ constitute a simultaneous solution to equations 4–6 of (1), then $A_1 = \tilde{Y}_1$, $A_2 = \tilde{Y}_2$ and $A_3 = \tilde{Y}_3$ constitute another solution if and only if*

$$\tilde{Y}_1 = T_4^{-1}Y_1T_6, \quad \tilde{Y}_2 = T_5^{-1}Y_2T_4 \quad \text{and} \quad \tilde{Y}_3 = T_6^{-1}Y_3T_5, \quad (15)$$

in which $T_i \in \mathcal{C}(B_i)$, $i = 4, 5, 6$.

Proof. We only prove (a); part (b) is similar. If X_1, X_2 and X_3 form a solution to equations 1–3 and \tilde{X}_1, \tilde{X}_2 and \tilde{X}_3 form another one, we can define

$$T_1 = I, \quad T_2 = X_2 X_3 \tilde{X}_3^{-1} \tilde{X}_2^{-1} \quad \text{and} \quad T_3 = X_3 \tilde{X}_3^{-1}.$$

It is easy to see that these matrices satisfy the equalities (14), and $T_i \in \mathcal{C}(B_i)$.

On the other hand, a simple calculation shows that if the matrices \tilde{X}_i satisfy (14), X_i being a solution of (1), then they also are a solution of (1). \square

The analog of Lemma 3 for n -cycles of square nonsingular matrices is also valid.

Now, there is a solution of the system (1) if and only if there is an intersection among the triples that solve 1–3 and those that solve 4–6. If $A_i = X_i, i = 1, 2, 3$, is a particular solution to 1–3 and $A_i = Y_i, i = 1, 2, 3$ is a particular solution to 4–6, then there is a common solution to 1–6 if and only if there exist $T_i \in \mathcal{C}(B_i), i = 1, \dots, 6$, such that

$$\tilde{X}_1 = \tilde{Y}_1, \quad \tilde{X}_2 = \tilde{Y}_2 \quad \text{and} \quad \tilde{X}_3 = \tilde{Y}_3,$$

in which $\tilde{X}_i, \tilde{Y}_i, i = 1, 2, 3$, are defined by (14) and (15). Letting $S_i = T_i^{-1}, i = 4, 5, 6$, algebraic manipulation shows that to solve the system (1) is equivalent to solving the system

$$T_1 S_4 Y_1 = X_1 T_2 S_6, \quad T_2 S_5 Y_2 = X_2 T_3 S_4, \quad T_3 S_6 Y_3 = X_3 T_1 S_5, \quad (16)$$

in which $T_i \in \mathcal{C}(B_i), i = 1, 2, 3$, and $S_i \in \mathcal{C}(B_i), i = 4, 5, 6$, are unknown matrices. We return to this later to understand the solvability of system (1) in a particular case.

We now know that a common Jordan form (i.e., mutual similarity) of the relevant B_i 's characterizes the solvability of a cycle of equations (e.g., 1–3 or 4–6). This generalizes in a simple way the square invertible case for $k = 2$ and may be generalized to the square invertible case for $k > 3$. What then about two of the equations, one from 1–3 and the other from 4–6. Each such pair amounts to transposition of the A indices. Somewhat surprisingly, condition (9) (for the relevant pair i, j) is sufficient (as well as necessary) for any particular pair of such equations. This may be seen using the observation of Sourour, which may be found in [5, p. 289].

Theorem 4. Suppose that $A, B \in M_n(\mathbb{C})$ and that $\det A \neq 0$. Then, for each non-cyclic permutation $\pi \neq I$, there exist $X_1, X_2, X_3 \in M_n(\mathbb{C})$ such that $A = X_1 X_2 X_3$ and $B = X_{\pi(1)} X_{\pi(2)} X_{\pi(3)}$ if and only if $\det B = \det A$.

Proof. The necessity of the condition follows from the multiplicativity of the determinant, as mentioned in (9). For sufficiency, there are two cases to consider (a) one in which π is the transposition $\pi(3) = 1, \pi(2) = 2$ and $\pi(1) = 3$ and (b) the other in which π is an adjacent transposition. Since $\det A = \det B \neq 0, \det AB^{-1} = 1$, we may write $AB^{-1} = PQ$ [5, p. 289] in which P has the distinct nonzero eigenvalues $\alpha_1, \alpha_2, \dots, \alpha_n$ and Q has the eigenvalues $1/\alpha_1, 1/\alpha_2, \dots, 1/\alpha_n$. Then Q^{-1} is similar to P , so that we may write $S^{-1}PS = Q^{-1}$ for some $S \in M_n(\mathbb{C})$. Letting

$C = B^{-1}Q^{-1}$, we have $AC = P$ and $SB = SQ^{-1}C^{-1} = PSC^{-1} = AC(SC^{-1})$. Thus $A = (SB)(CS^{-1})C^{-1}$ and $B = C^{-1}(CS^{-1})(SB)$, which completes the proof in case (a). Case (b) may be verified by a modification of the same argument, but we mention another argument. Again, since $\det AB^{-1} = 1$, we can write $AB^{-1} = XYX^{-1}Y^{-1}$ for some $X, Y \in M_n(\mathbb{C})$, is a multiplicative commutator if and only if its determinant is equal to 1 [5, p. 291]. Then $Y^{-1}X^{-1}A = X^{-1}Y^{-1}B =$ some matrix Z , so that $A = XYZ$ and $B = YXZ$, as was to be shown; the other adjacent transposition is similar. \square

It is not difficult to see that there is an exactly analogous result for a longer string of X 's and a single transposition. Theorem 4 lies in contrast to Flanders' observation for a transposition when there are only two matrices X_1, X_2 (so that a transposition is also a cycle). Then the much stronger condition of similarity, in place of equal determinants, is required.

We record as a theorem for comparison the case in which the permutation π is a cycle.

Theorem 5. *The system of equations 1–3 has a solution if and only if B_1, B_2 and B_3 are mutually similar (i.e., have a common Jordan form).*

This theorem is an immediate consequence of Theorem 3.1 of [3], because the matrices B_i are invertible.

Again, there is an analogous result for k equations corresponding to a k -cycle ($B_1 = A_1A_2 \cdots A_kA_1, B_2 = A_2 \cdots A_kA_1, \dots, B_k = A_kA_1 \cdots A_{k-1}$). For $k = 3$, Theorems 4 and 5 cover every possible pair of equations because every permutation is either a transposition or a cycle. For $k \geq 4$, this raises a natural question: what are the conditions on two (square, invertible) matrices A, B that they may be written $A = X_1X_2 \cdots X_k$ and $B = X_{\pi(1)}X_{\pi(2)} \cdots X_{\pi(k)}$, in which π is a permutation that is neither a cycle nor a transposition. Always, $\det A = \det B$ is necessary and A similar to B is sufficient.

Corollary 6. *Suppose that $A, B \in M_n(\mathbb{C})$ are invertible and π is a permutation of the set $\{1, 2, \dots, k\}$. Then A may be written $A = X_1X_2 \cdots X_k$ while B is written as $B = X_{\pi(1)}X_{\pi(2)} \cdots X_{\pi(k)}$, in which $X_i \in M_n(\mathbb{C}), i = 1, 2, \dots, k$, if and only if*

- (1) $B = A$ when π is the identity, or
- (2) B is similar to A when π is a cyclic permutation (that is, $\pi(1) = i, \pi(2) = i + 1, \dots, \pi(n) = i - 1$), or
- (3) $\det B = \det A$, otherwise.

We now turn to applying (16) to see more explicitly when the system (1) has a solution. Suppose that all of B_1, \dots, B_6 are diagonal with distinct nonzero eigenvalues. Then we know that

$$\begin{aligned}
B_1 &= \text{diag}(\lambda_1, \dots, \lambda_n), \\
B_2 &= \text{diag}(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)}), \\
B_3 &= \text{diag}(\lambda_{\tau(1)}, \dots, \lambda_{\tau(n)}), \\
B_4 &= \text{diag}(\mu_1, \dots, \mu_n), \\
B_5 &= \text{diag}(\mu_{\delta(1)}, \dots, \mu_{\delta(n)}), \\
B_6 &= \text{diag}(\mu_{\gamma(1)}, \dots, \mu_{\gamma(n)}),
\end{aligned} \tag{17}$$

in which σ , τ , δ and γ are permutations and

$$\prod_{i=1}^n \lambda_i = \prod_{i=1}^n \mu_i \tag{18}$$

(from equalities (8) and (9)). However, even in this special case, (8) and (9) do not insure a solution to (1).

Let P_σ (respectively, P_τ , P_δ , P_γ) be the permutation matrix corresponding to σ (respectively, τ , δ , γ); that is, $P_\sigma = [\delta_{\sigma(i),j}]$. Then, $P_\sigma B_1 P_\sigma^T = B_2$ ($P_\tau B_1 P_\tau^T = B_3$, $P_\delta B_4 P_\delta^T = B_5$, $P_\gamma B_4 P_\gamma^T = B_6$).

Up to invertible diagonal multiples, these matrices provide the unique similarities in (8) because $\mathcal{C}(B_i)$ is just the invertible diagonal matrices for each B_i . Now, with $X_1 = B_1 P_\sigma^T$, $X_2 = P_\sigma P_\tau^T$, $X_3 = P_\tau$, $Y_1 = B_4 P_\gamma^T$, $Y_2 = P_\delta$ and $Y_3 = P_\gamma P_\delta^T$ (from (12) and (13)), we have, from (16), that Y_1 and X_1 , Y_2 and X_2 , and Y_3 and X_3 differ only by invertible diagonal multiplication, so that each pair has common support (zero–nonzero pattern). It follows that

$$P_\sigma = P_\gamma, \quad P_\sigma = P_\delta P_\tau \quad \text{and} \quad P_\gamma = P_\tau P_\delta, \tag{19}$$

in order for a solution to exist.

In this case, we have the next important result.

Lemma 7. *Let B_1 , B_2 , B_3 , B_4 , B_5 and B_6 be defined as in (17), and assume that the conditions (19) hold. If (X_1, X_2, X_3) is a solution to equations 1–4 of the matrix system (1), then it is already a solution to the complete system.*

Proof. We know that $(B_1 P_\sigma^T, P_\delta, P_\tau)$ is a solution to equations 1–3 of (1). So, according to Lemma 3, all solutions of these three equations have the form

$$X_1 = T_1^{-1} B_1 P_\sigma^T T_2, \quad X_2 = T_2^{-1} P_\delta T_3 \quad \text{and} \quad X_3 = T_3^{-1} P_\tau T_1, \tag{20}$$

in which T_1 , T_2 and T_3 are nonsingular diagonal matrices. Clearly, $T_i \in \mathcal{C}(B_j)$ for all $i = 1, 2, 3$ and all $j = 1, 2, \dots, 6$.

Since (X_1, X_2, X_3) is also a solution to equation 4 of (1), we may write $X_1 X_3 X_2 = B_4$. Then,

$$\begin{aligned}
X_2 X_1 X_3 &= X_2 B_4 X_2^{-1} = T_2^{-1} P_\delta T_3 B_4 T_3^{-1} P_\delta^T T_2 \\
&= T_2^{-1} P_\delta B_4 P_\delta^T T_2 = T_2^{-1} B_5 T_2 = B_5
\end{aligned}$$

and

$$\begin{aligned} X_3 X_2 X_1 &= X_1^{-1} B_4 X_1 = T_2^{-1} P_\sigma B_1^{-1} T_1 B_4 T_1^{-1} B_1 P_\sigma^T T_2 \\ &= T_2^{-1} P_\sigma B_4 P_\sigma^T T_2 = T_2^{-1} B_6 T_2 = B_6, \end{aligned}$$

which completes the proof. \square

Conditions (19) may be restated as

$$P_\gamma = P_\sigma, \quad P_\sigma P_\delta = P_\delta P_\sigma \quad \text{and} \quad P_\tau = P_\sigma P_\delta^T, \quad (21)$$

a commutativity condition, together with a composition requirement. Thus, the subgroup of \mathcal{S}_n generated by $P_\sigma, P_\gamma, P_\delta, P_\tau$ is abelian and is generated by two of them, say P_σ, P_δ (other pairs suffice). If this subgroup is decomposable, there will be further conditions on $\lambda_1, \dots, \lambda_n$ and μ_1, \dots, μ_n that refine (18) in order that (1) be solvable; if not, (17), (18) and (21) will be sufficient. In the former case, the additional conditions will be of the form

$$\prod \lambda_i = \prod \mu_i,$$

in which each product is taken over indices i corresponding to an indecomposable component of the group generated by P_σ and P_δ .

Proposition 8. *Let B_1, B_2, B_3, B_4, B_5 and B_6 be defined as in (17), and assume that the conditions (19) hold. If $P_\sigma = P_{\sigma,1} \oplus P_{\sigma,2}$ and $P_\delta = P_{\delta,1} \oplus P_{\delta,2}$, where $P_{\sigma,i}$ and $P_{\delta,i}$ are n_i -by- n_i matrices ($i = 1, 2$), then*

$$\lambda_1 \cdots \lambda_{n_1} = \mu_1 \cdots \mu_{n_1} \quad \text{and} \quad \lambda_{n_1+1} \cdots \lambda_{n_1+n_2} = \mu_{n_1+1} \cdots \mu_{n_1+n_2}$$

are necessary for the system (1) to be solvable.

Proof. Suppose that (1) is solvable. Using (20), we have

$$\begin{aligned} B_4 &= (T_1^{-1} B_1 P_\sigma^T T_2) (T_3^{-1} P_\tau T_1) (T_2^{-1} P_\delta T_3) \\ &= T_1^{-1} B_1 P_\sigma^T T_2 T_3^{-1} P_\sigma P_\delta^T T_1 T_2^{-1} P_\delta T_3. \end{aligned} \quad (22)$$

If we consider the partitions (remember that T_i and B_j are diagonal matrices)

$$T_h = T_{h,1} \oplus T_{h,2}, \quad h = 1, 2, 3 \quad \text{and} \quad B_j = B_{j,1} \oplus B_{j,2}, \quad j = 1, \dots, 6,$$

in which $T_{h,i}$ and $B_{j,i}$ are n_i -by- n_i matrices ($i = 1, 2$), the equality (22) splits into two equalities, corresponding to the blocks $B_{4,1}$ and $B_{4,2}$. From these equalities, it is straightforward to see that $\det B_{4,i} = \det B_{1,i}$, $i = 1, 2$. \square

In order to be explicit about the above remarks, we need to study when two permutations commute and, in particular, a joint decomposition involving cycles when

they do. We say that an n -by- n permutation matrix is irreducible if it corresponds to the permutation

$$i_1 \rightarrow i_2 \rightarrow \cdots \rightarrow i_n \rightarrow i_1$$

for i_1, i_2, \dots, i_n some permutation of $1, 2, \dots, n$. A cyclic permutation is a special case, and any irreducible permutation matrix is a permutation similarity of the standard full cycle: $i_j = j$, $j = 1, 2, \dots, n$ (which we typically assume it is, without loss of generality). First, a lemma is useful.

Lemma 9. *Let Q be a $(0, 1)$ -matrix which has at most one element equal to 1 in each row and in each column. If Q commutes with an irreducible permutation P , then either Q is a power of P or $Q = O$.*

Proof. It is well-known that any irreducible permutation matrix is nonderogatory. In this case, it is also known (see, for example, [4, p. 135]) that a matrix Q commutes with P if and only if the matrix Q is a polynomial in P of degree at most $n - 1$, where n is the order of P ,

$$Q = a_0 I + a_1 P + \cdots + a_{n-1} P^{n-1}. \quad (23)$$

The structure of the matrices $Q, I, P, P^2, \dots, P^{n-1}$ imposes that there is at most one nonzero coefficient a_i in (23); moreover, the nonzero coefficient, if it exists, must be 1. \square

Now, we may characterize a commuting pair of permutation matrices by showing that they may be put in a common block form in which commutation is obvious. The blocks of this form will correspond to the refined product conditions on the λ 's and μ 's and, in turn, will lead to a characterization of the solvability of (1) for the problem in which the B_i 's are as in (17).

Theorem 10. *Let P and Q be n -by- n permutation matrices. Then P and Q commute if and only if, up to simultaneous permutation similarity,*

$$P = P_1 \oplus \cdots \oplus P_k \quad \text{and} \quad Q = Q_1 \oplus \cdots \oplus Q_k, \quad (24)$$

in which P_i and Q_i are n_i -by- n_i permutation matrices ($i = 1, \dots, k$; $n = \sum_{i=1}^k n_i$) and, for each i , one of the following possibilities holds:

- (a) P_i is a standard full cycle and Q_i is a power of P_i ;
- (b) $P_i = C_i \oplus \cdots \oplus C_i$ and

$$Q_i = \begin{bmatrix} O & I & & \\ \vdots & & \ddots & \\ \vdots & & & I \\ C_i^{q_i} & \cdots & \cdots & O \end{bmatrix},$$

in which C_i is a standard full cycle and q_i is a positive integer.

Proof. It is easy to see that if P and Q have the form (24) and each pair of matrices (P_i, Q_i) satisfies either (a) or (b), then P and Q commute. We will prove the converse implication.

Let T_1 be a permutation matrix such that $\tilde{P}_1 = T_1 P T_1^T = C \oplus \hat{P}_1$, where C is a standard full cycle. Let p be the order of C (we can identify C as having the greatest order among the standard full cycles in which the permutation represented by P decomposes, but it is not necessary). Let

$$\tilde{Q}_1 = T_1 Q T_1^T = \begin{bmatrix} Q_{11}^{(1)} & Q_{12}^{(1)} \\ Q_{21}^{(1)} & Q_{22}^{(1)} \end{bmatrix},$$

in which $Q_{11}^{(1)}$ has the same size as C .

The commutation condition, $\tilde{P}_1 \tilde{Q}_1 = \tilde{Q}_1 \tilde{P}_1$, implies that

$$C Q_{11}^{(1)} = Q_{11}^{(1)} C \quad \text{and} \quad \hat{P}_1 Q_{22}^{(1)} = Q_{22}^{(1)} \hat{P}_1. \quad (25)$$

According to Lemma 9, we know that either $Q_{11}^{(1)} = C^q$, for some $q \in \mathbb{Z}^+$, or $Q_{11}^{(1)} = O$.

If $Q_{11}^{(1)} = C^q$, we take

$$P_1 = C \quad \text{and} \quad Q_1 = C^q. \quad (26)$$

In this case, it is clear that $Q_{12}^{(1)}$ and $Q_{21}^{(1)}$ must be null matrices. Then, the matrices $P'_1 = \hat{P}_1$ and $Q'_1 = Q_{22}^{(1)}$ satisfy $\tilde{P}_1 = P_1 \oplus P'_1$, $\tilde{Q}_1 = Q_1 \oplus Q'_1$. Moreover, the size of the matrices P'_1 and Q'_1 is smaller than the size of P and Q and, according to (25), these new matrices are still commuting.

Otherwise, if $Q_{11}^{(1)} = O$, it is possible to permute the columns of the matrix \tilde{Q}_1 in order to obtain that the first p columns corresponding to the block $Q_{12}^{(1)}$ form the identity matrix. Observe that the permutation which is used in this reordering can be chosen so that it leaves the first p indices fixed. Thus, there exists a permutation matrix T_2 such that

$$\tilde{P}_2 = T_2 \tilde{P}_1 T_2^T = \begin{bmatrix} C & O & O \\ O & P_{22}^{(2)} & P_{23}^{(2)} \\ O & P_{32}^{(2)} & P_{33}^{(2)} \end{bmatrix}$$

and

$$\tilde{Q}_2 = T_2 \tilde{Q}_1 T_2^T = \begin{bmatrix} O & I & O \\ Q_{21}^{(2)} & O & Q_{23}^{(2)} \\ Q_{31}^{(2)} & O & Q_{33}^{(2)} \end{bmatrix}.$$

The equality $\tilde{P}_2 \tilde{Q}_2 = \tilde{Q}_2 \tilde{P}_2$ implies that

$$P_{22}^{(2)} = C, \quad C Q_{21}^{(2)} = Q_{21}^{(2)} C \quad \text{and} \quad P_{33}^{(2)} Q_{33}^{(2)} = Q_{33}^{(2)} P_{33}^{(2)}.$$

Since \tilde{P}_2 is a permutation matrix, the first of these equalities provides that $P_{23}^{(2)} = O$ and $P_{32}^{(2)} = O$. Moreover, the second of them assures that either $Q_{21}^{(2)} = C^q$ or $Q_{21}^{(2)} = O$. In the first case, we get

$$P_1 = C \oplus C \quad \text{and} \quad Q_1 = \begin{bmatrix} O & I \\ C^q & O \end{bmatrix}.$$

If $Q_{21}^{(2)} = O$, we repeat the process until we find a block $Q_{r1}^{(r)} \neq O$. It is clear that we will be able to find this nonnull block, since the successive matrices \tilde{Q}_r are permutation (and finite) matrices. Thus, we will get

$$P_1 = \underbrace{C \oplus \cdots \oplus C}_r, \quad Q_1 = \begin{bmatrix} O & I & & \\ \vdots & & \ddots & \\ \vdots & & & I \\ C^q & \cdots & \cdots & O \end{bmatrix} \quad (27)$$

and P and Q are simultaneously permutation similar to $P_1 \oplus P'_1$ and $Q_1 \oplus Q'_1$, respectively, where the matrices P'_1 and Q'_1 commute and have smaller size than that of P and Q .

Obviously, in every event, the matrices P and Q , up to simultaneous permutation similarity, can be written as

$$P = P_1 \oplus P'_1, \quad Q = Q_1 \oplus Q'_1,$$

in which P_1, Q_1 are given either by (26) or (27) and the matrices P'_1 and Q'_1 commute. Then, we can repeat the above process on the pair of permutation matrices P'_1 and Q'_1 . Since each step of this process decreases the size of the considered matrices, the procedure is finite and completes the proof. \square

Given a commuting pair of permutation matrices P and Q , according to Theorem 10, there is a natural finest partition $N_1 \cup N_2 \cup \cdots \cup N_k = N$ of $N = \{1, 2, \dots, n\}$ in which N_i is the set of indices corresponding to the blocks P_i and Q_i prior to the permutation of P and Q into the special form given in Theorem 10. We call this the *commutation partition* of P and Q , and, according to Proposition 8, this partition corresponds to the conditions on the λ 's and μ 's needed to solve problem (1). Notice that if either P_σ or P_δ is irreducible, the commutation partition is simply N , in which case the only condition will be (18), which is the successor to (9) in the case (17). To illustrate what can occur, we first give three examples.

Example 11. Let

$$P_\sigma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad P_\delta = P_\sigma^2.$$

Observe that the permutation P_σ is a standard full cycle ($\sigma: 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$). According to (17) and (21), we have

$$\begin{aligned} B_1 &= \text{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4), \\ B_2 &= \text{diag}(\lambda_2, \lambda_3, \lambda_4, \lambda_1), \\ B_3 &= \text{diag}(\lambda_4, \lambda_1, \lambda_2, \lambda_3), \\ B_4 &= \text{diag}(\mu_1, \mu_2, \mu_3, \mu_4), \\ B_5 &= \text{diag}(\mu_3, \mu_4, \mu_1, \mu_2), \\ B_6 &= \text{diag}(\mu_2, \mu_3, \mu_4, \mu_1). \end{aligned}$$

We know that all solutions to equations 1–3 of (1) are given by (20). Let $T_i = \text{diag}(t_{i1}, t_{i2}, t_{i3}, t_{i4})$, for $i = 1, 2, 3$.

The matrix equation $X_1 X_3 X_2 = B_4$ can be written as

$$\frac{t_{13}t_{24}t_{31}}{t_{11}t_{23}t_{34}} = \frac{\mu_1}{\lambda_1}, \quad \frac{t_{14}t_{21}t_{32}}{t_{12}t_{24}t_{31}} = \frac{\mu_2}{\lambda_2}, \quad \frac{t_{11}t_{22}t_{33}}{t_{13}t_{21}t_{32}} = \frac{\mu_3}{\lambda_3}, \quad \frac{t_{12}t_{23}t_{34}}{t_{14}t_{22}t_{33}} = \frac{\mu_4}{\lambda_4}. \quad (28)$$

A simple calculation shows that the equality $\lambda_1\lambda_2\lambda_3\lambda_4 = \mu_1\mu_2\mu_3\mu_4$ is the unique condition so that the system (28) be solvable. To obtain a solution to this system, we may take $T_1 = T_2 = I$ and

$$T_3 = \text{diag}\left(\frac{\mu_1}{\lambda_1}, \frac{\mu_1\mu_2}{\lambda_1\lambda_2}, \frac{\mu_1\mu_2\mu_3}{\lambda_1\lambda_2\lambda_3}, 1\right).$$

Hence, the matrices

$$X_1 = \begin{bmatrix} 0 & 0 & 0 & \lambda_1 \\ \lambda_2 & 0 & 0 & 0 \\ 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & \lambda_4 & 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 & 0 & \frac{\mu_1\mu_2\mu_3}{\lambda_1\lambda_2\lambda_3} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{\mu_1}{\lambda_1} & 0 & 0 & 0 \\ 0 & \frac{\mu_1\mu_2}{\lambda_1\lambda_2} & 0 & 0 \end{bmatrix}$$

and

$$X_3 = \begin{bmatrix} 0 & 0 & 0 & \frac{\lambda_1}{\mu_1} \\ \frac{\lambda_1\lambda_2}{\mu_1\mu_2} & 0 & 0 & 0 \\ 0 & \frac{\lambda_1\lambda_2\lambda_3}{\mu_1\mu_2\mu_3} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

form a solution to equations 1–4 of (1). Lemma 7 assures that this solution is already a solution to the complete matrix system (1).

Example 12. Let

$$P_\sigma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad P_\delta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Observe that the permutation matrices P_σ and P_δ are decomposable: $P_\sigma = C \oplus C$ and $P_\delta = C^2 \oplus C$, where C is the standard full cycle of order 2. In this case, we have

$$B_1 = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4),$$

$$B_2 = \text{diag}(\lambda_2, \lambda_1, \lambda_4, \lambda_3),$$

$$B_3 = \text{diag}(\lambda_2, \lambda_1, \lambda_3, \lambda_4),$$

$$B_4 = \text{diag}(\mu_1, \mu_2, \mu_3, \mu_4),$$

$$B_5 = \text{diag}(\mu_1, \mu_2, \mu_4, \mu_3),$$

$$B_6 = \text{diag}(\mu_2, \mu_1, \mu_4, \mu_3).$$

Again, all solutions to equations 1–3 are given by (20). Now, the matrix equation $X_1 X_3 X_2 = B_4$ implies that

$$\frac{t_{22}t_{31}}{t_{21}t_{32}} = \frac{\mu_1}{\lambda_1}, \quad \frac{t_{21}t_{32}}{t_{22}t_{31}} = \frac{\mu_2}{\lambda_2}, \quad \frac{t_{14}t_{33}}{t_{13}t_{34}} = \frac{\mu_3}{\lambda_3}, \quad \frac{t_{13}t_{34}}{t_{14}t_{33}} = \frac{\mu_4}{\lambda_4}.$$

Notice that to solve this system, the conditions $\lambda_1\lambda_2 = \mu_1\mu_2$ and $\lambda_3\lambda_4 = \mu_3\mu_4$ are necessary. The matrices $T_1 = T_2 = I$ and

$$T_3 = \text{diag}\left(\frac{\mu_1}{\lambda_1}, 1, \frac{\mu_3}{\lambda_3}, 1\right)$$

provide a solution to the matrix system:

$$X_1 = \begin{bmatrix} 0 & \lambda_1 & 0 & 0 \\ \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_3 \\ 0 & 0 & \lambda_4 & 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} \frac{\mu_1}{\lambda_1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{\mu_3}{\lambda_3} & 0 \end{bmatrix}$$

and

$$X_3 = \begin{bmatrix} 0 & \frac{\lambda_1}{\mu_1} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{\lambda_3}{\mu_3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Example 13. Finally, let

$$P_\sigma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad P_\delta = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

(these matrices have the structure given in (27), where C is the standard full cycle of order 2 and $q = 2$). Then,

$$\begin{aligned} B_1 &= \text{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4), \\ B_2 &= \text{diag}(\lambda_2, \lambda_1, \lambda_4, \lambda_3), \\ B_3 &= \text{diag}(\lambda_4, \lambda_3, \lambda_2, \lambda_1), \\ B_4 &= \text{diag}(\mu_1, \mu_2, \mu_3, \mu_4), \\ B_5 &= \text{diag}(\mu_3, \mu_4, \mu_1, \mu_2), \\ B_6 &= \text{diag}(\mu_2, \mu_1, \mu_4, \mu_3). \end{aligned}$$

The matrix equation $X_1 X_3 X_2 = B_4$ can be written as

$$\frac{t_{13}t_{22}t_{31}}{t_{11}t_{23}t_{32}} = \frac{\mu_1}{\lambda_1}, \quad \frac{t_{14}t_{21}t_{32}}{t_{12}t_{24}t_{31}} = \frac{\mu_2}{\lambda_2}, \quad \frac{t_{11}t_{24}t_{33}}{t_{13}t_{21}t_{34}} = \frac{\mu_3}{\lambda_3}, \quad \frac{t_{12}t_{23}t_{34}}{t_{14}t_{22}t_{33}} = \frac{\mu_4}{\lambda_4}.$$

This system is solvable if and only if $\lambda_1\lambda_2\lambda_3\lambda_4 = \mu_1\mu_2\mu_3\mu_4$. A solution to it is given by

$$T_1 = \text{diag}\left(1, \frac{\lambda_1\lambda_2}{\mu_1\mu_2}, 1, 1\right), \quad T_2 = Id$$

and

$$T_3 = \text{diag}\left(1, \frac{\lambda_1}{\mu_1}, \frac{\lambda_1\lambda_2\lambda_4}{\mu_1\mu_2\mu_4}, 1\right).$$

Then, the matrices

$$\begin{aligned} X_1 &= \begin{bmatrix} 0 & \lambda_1 & 0 & 0 \\ \frac{\mu_1\mu_2}{\lambda_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_3 \\ 0 & 0 & \lambda_4 & 0 \end{bmatrix}, & X_2 &= \begin{bmatrix} 0 & 0 & \frac{\lambda_1\lambda_2\lambda_4}{\mu_1\mu_2\mu_4} & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{\lambda_1}{\mu_1} & 0 & 0 \end{bmatrix}, \\ X_3 &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{\mu_1}{\lambda_1} & 0 \\ 0 & \frac{\mu_4}{\lambda_4} & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

form a solution to the matrix system (1).

Our aim is to find a necessary and sufficient condition so that the system (1) be solvable when the matrices B_i are as in (17). We know that conditions (19) are necessary. From these conditions, the permutation matrices P_σ and P_δ commute. Then, according to Theorem 10, there exists a permutation similarity S such that

$$S^{-1}P_\sigma S = \widehat{P}_\sigma = P_{\sigma,1} \oplus \cdots \oplus P_{\sigma,k} \quad \text{and} \quad S^{-1}P_\delta S = \widehat{P}_\delta = P_{\delta,1} \oplus \cdots \oplus P_{\delta,k},$$

where $P_{\sigma,i}$ and $P_{\delta,i}$ are n_i -by- n_i permutation matrices, for $i = 1, \dots, k$, and each pair $(P_{\sigma,i}, P_{\delta,i})$ has either the form (C, C^q) or the form given in (27); in both cases, C is a standard full cycle and q is a positive integer.

If we define $\widehat{B}_i = S^{-1}B_iS$, for $i = 1, \dots, 6$, and $(\widehat{X}_1, \widehat{X}_2, \widehat{X}_3)$ is a solution to the system

$$\begin{aligned}\widehat{B}_1 &= \widehat{X}_1 \widehat{X}_2 \widehat{X}_3, & \widehat{B}_2 &= \widehat{X}_2 \widehat{X}_3 \widehat{X}_1, & \widehat{B}_3 &= \widehat{X}_3 \widehat{X}_1 \widehat{X}_2, \\ \widehat{B}_4 &= \widehat{X}_1 \widehat{X}_3 \widehat{X}_2, & \widehat{B}_5 &= \widehat{X}_2 \widehat{X}_1 \widehat{X}_3, & \widehat{B}_6 &= \widehat{X}_3 \widehat{X}_2 \widehat{X}_1,\end{aligned}$$

then $X_1 = S\widehat{X}_1S^{-1}$, $X_2 = S\widehat{X}_2S^{-1}$ and $X_3 = S\widehat{X}_3S^{-1}$ form a solution to (1). Moreover, notice that $B_2 = P_{\sigma}B_1P_{\sigma}^T$ if and only if $\widehat{B}_2 = \widehat{P}_{\sigma}\widehat{B}_1\widehat{P}_{\sigma}^T$, and so on for the rest of the known relations involving the B_i 's. Therefore, without loss of generality, we may suppose that $P_{\sigma} = \widehat{P}_{\sigma}$ and $P_{\delta} = \widehat{P}_{\delta}$. Then the system (1) splits into the subsystems

$$\begin{aligned}B_{1,i} &= X_{1,i}X_{2,i}X_{3,i}, & B_{2,i} &= X_{2,i}X_{3,i}X_{1,i}, & B_{3,i} &= X_{3,i}X_{1,i}X_{2,i}, \\ B_{4,i} &= X_{1,i}X_{3,i}X_{2,i}, & B_{5,i} &= X_{2,i}X_{1,i}X_{3,i}, & B_{6,i} &= X_{3,i}X_{2,i}X_{1,i}\end{aligned}$$

for $i = 1, \dots, k$, in which $B_{2,i} = P_{\sigma,i}B_{1,i}P_{\sigma,i}^T$, $B_{3,i} = P_{\tau,i}B_{1,i}P_{\tau,i}^T$, $B_{5,i} = P_{\delta,i}B_{4,i}P_{\delta,i}^T$ and $B_{6,i} = P_{\sigma,i}B_{4,i}P_{\sigma,i}^T$.

So, we only need to study the solvability of (1) in two cases:

- (1) P_{σ} is a standard full cycle and P_{δ} is a power of P_{σ} , and
- (2) the pair (P_{σ}, P_{δ}) has the structure given in (27).

Notice that (22) can also be written, after some algebraic manipulations, as

$$B_4B_1^{-1} = T_1^{-1}(P_{\delta}^T T_1 P_{\delta})(P_{\sigma}^T T_2 P_{\sigma})(P_{\delta}^T T_2^{-1} P_{\delta})(P_{\sigma}^T T_3^{-1} P_{\sigma})T_3.$$

From Examples 11–13, it appears possible to consider $T_2 = I$; we will test this idea. Under this supposition, the equation to solve is

$$B_4B_1^{-1} = T_1^{-1}(P_{\delta}^T T_1 P_{\delta})(P_{\sigma}^T T_3^{-1} P_{\sigma})T_3,$$

which is equivalent to

$$\frac{t_{1,\delta^{-1}(i)} \cdot t_{3,i}}{t_{1,i} \cdot t_{3,\sigma^{-1}(i)}} = \frac{\mu_i}{\lambda_i} \quad \text{for } i = 1, \dots, n. \quad (29)$$

Proposition 14. Let B_1, B_2, B_3, B_4, B_5 and B_6 be defined as in (17) and assume that the conditions (19) hold. Assume also that P_{σ} corresponds to a standard full cycle of order n and P_{δ} is a power of P_{σ} . Then, the system (1) is solvable if and only if $\det B_1 = \det B_4$; moreover, a solution to the system is given by

$$X_1 = B_1 P_{\sigma}^T, \quad X_2 = P_{\delta} T_3 \quad \text{and} \quad X_3 = T_3^{-1} P_{\tau}, \quad (30)$$

where

$$T_3 = \text{diag} \left(\frac{\mu_1}{\lambda_1}, \frac{\mu_1 \mu_2}{\lambda_1 \lambda_2}, \dots, \frac{\mu_1 \cdots \mu_{n-1}}{\lambda_1 \cdots \lambda_{n-1}}, 1 \right). \quad (31)$$

Proof. The condition $\det B_1 = \det B_4$ is obviously necessary. Moreover, using the matrices defined by (30), it is straightforward to see that

$$X_1 X_2 X_3 = B_1, \quad X_2 X_3 X_1 = B_2 \quad \text{and} \quad X_3 X_1 X_2 = B_3,$$

and

$$X_1 X_3 X_2 = B_1 P_\sigma^T T_3^{-1} P_\tau P_\delta T_3 = B_1 P_\sigma^T T_3^{-1} P_\sigma T_3. \quad (32)$$

Then, the definition of T_3 given in (31) and the special form of P_σ (that is, $\sigma(i) = i + 1$ for $i = 1, \dots, n - 1$ and $\sigma(n) = 1$) provide

$$P_\sigma^T T_3^{-1} P_\sigma = \text{diag} \left(1, \frac{\lambda_1}{\mu_1}, \dots, \frac{\lambda_1 \cdots \lambda_{n-2}}{\mu_1 \cdots \mu_{n-2}}, \frac{\lambda_1 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_{n-1}} \right). \quad (33)$$

From (17) and (31)–(33),

$$\begin{aligned} X_1 X_3 X_2 &= \text{diag} \left(\mu_1, \mu_2, \dots, \mu_{n-1}, \frac{\lambda_1 \cdots \lambda_n}{\mu_1 \cdots \mu_{n-1}} \right) \\ &= \text{diag} \left(\mu_1, \mu_2, \dots, \mu_{n-1}, \frac{\mu_n \det B_1}{\det B_4} \right). \end{aligned}$$

Then, $\det B_1 = \det B_4$ is a sufficient condition so that $X_1 X_3 X_2 = B_4$. Lemma 7 completes the proof. \square

To find a solution to system (1) in the case in which P_σ and P_δ have the structure given in (27), we need some preliminary results.

As usual, we denote by $\lfloor x \rfloor$ the integer part of x and by $\text{mod}(x, y)$ the remainder of x after division by y .

Lemma 15. Let $P_\sigma = \underbrace{C \oplus \cdots \oplus C}_r$, in which C is a standard full cycle of order p and let $1 \leq i \leq pr$. Then,

- (1) $i = \lfloor \frac{i-1}{p} \rfloor p + \text{mod}(i-1, p) + 1$,
- (2) $\sigma^{-1}(i) = \lfloor \frac{i-1}{p} \rfloor p + \text{mod}(i-2, p) + 1$.

Proof. Part (1) is straightforward to prove:

$$i = (i-1) + 1 = \left\lfloor \frac{i-1}{p} \right\rfloor p + \text{mod}(i-1, p) + 1.$$

To prove part (2), observe that the permutation σ may be described as follows:

$$\begin{aligned} 1 &\rightarrow 2 \rightarrow \cdots \rightarrow p \rightarrow 1, & (p+1) &\rightarrow (p+2) \cdots \rightarrow 2p \rightarrow (p+1), \\ \dots, & [(r-1)p+1] &\rightarrow [(r-1)p+2] \rightarrow \cdots \rightarrow rp &\rightarrow [(r-1)p+1]. \end{aligned}$$

We analyze three cases.

- If $i = ap + 1$ with $0 \leq a < r$, we have $\sigma^{-1}(i) = i + p - 1$ and

$$\left\lfloor \frac{i-1}{p} \right\rfloor p + \text{mod}(i-2, p) + 1 = ap + (p-1) + 1 = i + p - 1.$$

- If $i = ap + h$ with $0 \leq a < r$ and $2 \leq h \leq p-1$, we know that $\sigma^{-1}(i) = i-1$; on the other hand,

$$\left\lfloor \frac{i-1}{p} \right\rfloor p + \text{mod}(i-2, p) + 1 = ap + (h-2) + 1 = ap + h - 1 = i - 1.$$

- If $i = ap$ with $1 \leq a \leq r$, then $\sigma^{-1}(i) = i-1$ and

$$\begin{aligned} \left\lfloor \frac{i-1}{p} \right\rfloor p + \text{mod}(i-2, p) + 1 &= (a-1)p + (p-2) + 1 \\ &= ap - 1 = i - 1. \quad \square \end{aligned}$$

Lemma 16. If $p, r, i \in \mathbb{Z}^+$, we can write

$$i = (j-1)p + \text{mod}(j-k, p) + 1, \quad (34)$$

in which

$$j = \left\lfloor \frac{i-1}{p} \right\rfloor + 1 \quad \text{and} \quad k = \text{mod}\left(1 + \left\lfloor \frac{i-1}{p} \right\rfloor - i, p\right) + 1.$$

Proof. According to part (1) of Lemma 15, if we write i as proposed in (34), we have

$$j-1 = \left\lfloor \frac{i-1}{p} \right\rfloor \quad \text{and} \quad \text{mod}(j-k, p) = \text{mod}(i-1, p).$$

The second of these equalities implies that $\text{mod}(k-1, p) = \text{mod}(j-i, p)$, that is,

$$\text{mod}(k-1, p) = \text{mod}\left(1 + \left\lfloor \frac{i-1}{p} \right\rfloor - i, p\right).$$

In order that this relation be true, we may take $k = \text{mod}\left(1 + \left\lfloor \frac{i-1}{p} \right\rfloor - i, p\right) + 1$. \square

Fixing $p \in \mathbb{Z}^+$, we define the functions $J_p(i)$ and $K_p(i)$ as

$$J_p(i) = \left\lfloor \frac{i-1}{p} \right\rfloor + 1, \quad K_p(i) = \text{mod}\left(1 + \left\lfloor \frac{i-1}{p} \right\rfloor - i, p\right) + 1.$$

They are the unique values such that

$$\begin{aligned} i &= [J_p(i) - 1]p + \text{mod}[J_p(i) - K_p(i), p] + 1, \\ J_p(i) &\geq 1 \quad \text{and} \quad 1 \leq K_p(i) \leq p. \end{aligned} \quad (35)$$

It is easy to see that if $i = (j-1)p + \text{mod}(j-k, p) + 1$, then

$$\sigma^{-1}(i) = (j-1)p + \text{mod}(j-k-1, p) + 1.$$

In other words,

$$J_p(\sigma^{-1}(i)) = J_p(i) \quad \text{and} \quad K_p(\sigma^{-1}(i)) = \text{mod}(K_p(i), p) + 1.$$

Now, we need to study the permutation δ . Let $\hat{\delta}$ be the permutation corresponding to C^q . Then, δ can be described by

$$\begin{aligned} 1 &\rightarrow (p+1) \rightarrow (2p+1) \rightarrow \cdots \rightarrow [(r-1)p+1] \rightarrow \hat{\delta}(1), \\ 2 &\rightarrow (p+2) \rightarrow (2p+2) \rightarrow \cdots \rightarrow [(r-1)p+2] \rightarrow \hat{\delta}(2), \\ &\vdots \qquad \qquad \qquad \vdots \\ p &\rightarrow 2p \rightarrow 3p \rightarrow \cdots \rightarrow rp \rightarrow \hat{\delta}(p), \end{aligned}$$

where, obviously, $1 \leq \hat{\delta}(i) \leq p$, for $i = 1, 2, \dots, p$. From these chains of images we may construct the cycles of δ , according to the values $\hat{\delta}(i)$. It is easy to see that if $\hat{\delta}$ is an irreducible permutation, then δ is also irreducible; in this case, the solution of the matrix system studied can be constructed as in the case solved above, interchanging there the roles of T_1 and T_3 .

For $p < i \leq rp$, it is straightforward to prove that $J_p(\delta^{-1}(i)) = J_p(i) - 1$ and $K_p(\delta^{-1}(i)) = \text{mod}(K_p(i) - 2, p) + 1$.

To simplify the notation, we define the values, for $n \in \mathbb{Z}^+$ and $m = 1, 2, \dots, p$,

$$\alpha(n, m) = \frac{\lambda_{(n-1)p + \text{mod}(n-m, p) + 1}}{\mu_{(n-1)p + \text{mod}(n-m, p) + 1}}. \quad (36)$$

Lemma 17. *The values defined in (36) satisfy the following properties:*

- (1) $\alpha(J_p(i), K_p(i)) = \frac{\lambda_i}{\mu_i}$;
- (2) $\prod_{m=1}^p \alpha(n, m) = \prod_{m=(n-1)p+1}^{np} \frac{\lambda_m}{\mu_m}$.

Proof. Part (1) follows from (35). To prove part (2), notice that, for $m = 1, 2, \dots, p$,

$$(n-1)p + 1 \leq (n-1)p + \text{mod}(n-m, p) + 1 \leq np.$$

Since $\text{mod}(n-m_1, p) \neq \text{mod}(n-m_2, p)$ when $m_1 \neq m_2$, the expression given by

$$(n-1)p + \text{mod}(n-m, p) + 1$$

takes p distinct values, that is, *all* the integer values from $(n-1)p + 1$ to np inclusive. \square

Proposition 18. *Let B_1, B_2, B_3, B_4, B_5 and B_6 be defined as in (17) and assume that the conditions (19) hold. Assume also that P_σ and P_δ have the structure described in (27), in which C is a standard full cycle of order p . Then, the system (1)*

is solvable if and only if $\det B_1 = \det B_4$; moreover, a solution to the system is given by

$$X_1 = T_1^{-1} B_1 P_\sigma^T, \quad X_2 = P_\delta T_3 \quad \text{and} \quad X_3 = T_3^{-1} P_\tau T_1,$$

where $T_h = \text{diag}(t_{h,1}, t_{h,2}, \dots, t_{h,rp})$, $h = 1, 3$, and

$$t_{1,i} = \begin{cases} \prod_{s=1}^{J_p(i) \cdot p} \frac{\lambda_s}{\mu_s} & \text{if } K_p(i) = p \text{ and } 1 \leq J_p(i) \leq r-1, \\ 1 & \text{otherwise,} \end{cases} \quad (37)$$

$$t_{3,i} = \begin{cases} 1 & \text{if } K_p(i) = 1, \\ \left(\prod_{s=1}^{[J_p(i)-1]p} \frac{\lambda_s}{\mu_s} \right) \left(\prod_{m=1}^{K_p(i)-1} \alpha(J_p(i), m) \right) & \text{if } 2 \leq K_p(i) \leq p. \end{cases} \quad (38)$$

Proof. To prove this, we determine a sufficient condition so that the values given by (37) and (38) satisfy Eqs. (29).

(1) For $1 \leq i \leq p$, we know that $J_p(i) = 1$ and $(r-1)p + 1 \leq \delta^{-1}(i) \leq rp$. Since

$$J_p(\delta^{-1}(i)) = \left\lfloor \frac{\delta^{-1}(i) - 1}{p} \right\rfloor + 1 = r,$$

in the definition (37) we find that $t_{1,\delta^{-1}(i)} = 1$. We need still to study three sub-cases.

(a) If $K_p(i) = 1$, we know that $i = 1$. Since $K_p(\sigma^{-1}(1)) = 2$, according to (38), we have

$$\frac{t_{1,\delta^{-1}(1)} \cdot t_{3,1}}{t_{1,1} \cdot t_{3,\sigma^{-1}(1)}} = \frac{1 \cdot 1}{1 \cdot \alpha(J_p(1), 1)} = \frac{\mu_1}{\lambda_1}.$$

(b) For $2 \leq K_p(i) \leq p-1$ (that is, for $3 \leq i \leq p$), we know that $K_p(\sigma^{-1}(i)) = K_p(i) + 1$. Then,

$$\frac{t_{1,\delta^{-1}(i)} \cdot t_{3,i}}{t_{1,i} \cdot t_{3,\sigma^{-1}(i)}} = \frac{1 \cdot \prod_{m=1}^{K_p(i)-1} \alpha(J_p(i), m)}{1 \cdot \prod_{m=1}^{K_p(i)} \alpha(J_p(i), m)} = \frac{1}{\alpha(J_p(i), K_p(i))} = \frac{\mu_i}{\lambda_i}.$$

(c) If $K_p(i) = p$, that is, if $i = 2$, we have $K_p(\sigma^{-1}(i)) = 1$. Then,

$$\begin{aligned} \frac{t_{1,\delta^{-1}(2)} \cdot t_{3,2}}{t_{1,2} \cdot t_{3,\sigma^{-1}(2)}} &= \frac{1 \cdot \prod_{m=1}^{p-1} \alpha(J_p(2), m)}{\prod_{s=1}^p \frac{\lambda_s}{\mu_s} \cdot 1} = \frac{\prod_{m=1}^p \alpha(J_p(2), m)}{\alpha(J_p(2), p) \cdot \prod_{s=1}^p \frac{\lambda_s}{\mu_s}} \\ &= \frac{1}{\alpha(J_p(2), p)} = \frac{\mu_2}{\lambda_2}. \end{aligned}$$

- (2) For $p + 1 \leq i \leq rp$, we know that $2 \leq J_p(i) \leq r$ and $J_p(\delta^{-1}(i)) = J_p(i) - 1$. Then,

(a) If $K_p(i) = 1$, we have $K_p(\sigma^{-1}(i)) = 2$ and $K_p(\delta^{-1}(i)) = p$. Thus,

$$\frac{t_{1,\delta^{-1}(i)} \cdot t_{3,i}}{t_{1,i} \cdot t_{3,\sigma^{-1}(i)}} = \frac{\prod_{s=1}^{[J_p(i)-1]p} \frac{\lambda_s}{\mu_s} \cdot 1}{1 \cdot \left(\prod_{s=1}^{[J_p(i)-1]p} \frac{\lambda_s}{\mu_s} \right) \cdot \alpha(J_p(i), 1)} = \frac{\mu_i}{\lambda_i}.$$

- (b) For $2 \leq K_p(i) \leq p - 1$, we know that $K_p(\sigma^{-1}(i)) = K_p(i) + 1$ and $K_p(\delta^{-1}(i)) = K_p(i) - 1$. Then,

$$\begin{aligned} \frac{t_{1,\delta^{-1}(i)} \cdot t_{3,i}}{t_{1,i} \cdot t_{3,\sigma^{-1}(i)}} &= \frac{1 \cdot \left(\prod_{s=1}^{[J_p(i)-1]p} \frac{\lambda_s}{\mu_s} \right) \left(\prod_{m=1}^{K_p(i)-1} \alpha(J_p(i), m) \right)}{1 \cdot \left(\prod_{s=1}^{[J_p(i)-1]p} \frac{\lambda_s}{\mu_s} \right) \left(\prod_{m=1}^{K_p(i)} \alpha(J_p(i), m) \right)} \\ &= \frac{1}{\alpha(J_p(i), K_p(i))} = \frac{\mu_i}{\lambda_i}. \end{aligned}$$

- (c) Finally, we study the subcase in which $K_p(i) = p$. Since $K_p(\sigma^{-1}(i)) = 1$ and $K_p(\delta^{-1}(i)) = p - 1$, we can write

$$\begin{aligned} \frac{t_{1,\delta^{-1}(i)} \cdot t_{3,i}}{t_{1,i} \cdot t_{3,\sigma^{-1}(i)}} &= \frac{1 \cdot \left(\prod_{s=1}^{[J_p(i)-1]p} \frac{\lambda_s}{\mu_s} \right) \left(\prod_{m=1}^{p-1} \alpha(J_p(i), m) \right)}{t_{1,i} \cdot 1} \\ &= \frac{\prod_{s=1}^{J_p(i)} \frac{\lambda_s}{\mu_s}}{t_{1,i} \cdot \alpha(J_p(i), p)}. \end{aligned}$$

In this expression, we need yet to consider two cases: $J_p(i) < r$ and $J_p(i) = r$. When $J_p(i) < r$, we have

$$\frac{\prod_{s=1}^{J_p(i)} \frac{\lambda_s}{\mu_s}}{t_{1,i} \cdot \alpha(J_p(i), p)} = \frac{\prod_{s=1}^{J_p(i)} \frac{\lambda_s}{\mu_s}}{\left(\prod_{s=1}^{J_p(i)} \frac{\lambda_s}{\mu_s} \right) \alpha(J_p(i), p)} = \frac{1}{\alpha(J_p(i), K_p(i))} = \frac{\mu_i}{\lambda_i}.$$

On the other hand, when $J_p(i) = r$, we know that $t_{1,i} = 1$. Then,

$$\frac{t_{1,\delta^{-1}(i)} \cdot t_{3,i}}{t_{1,i} \cdot t_{3,\sigma^{-1}(i)}} = \frac{\prod_{s=1}^r \frac{\lambda_s}{\mu_s}}{\alpha(J_p(i), K_p(i))} = \frac{\det B_1}{\det B_4} \cdot \frac{\mu_i}{\lambda_i}.$$

This is the unique case in which we need an additional condition so that the corresponding equation given in (29) be satisfied: it must be that $\det B_1 = \det B_4$.

Hence, it is clear that the necessary condition $\det B_1 = \det B_4$ is also a sufficient condition for solvability of (1). \square

Using the results found above, we have proved our final main theorem.

Theorem 19. *Let B_1, \dots, B_6 be as in (17), with $\lambda_1, \dots, \lambda_n$ distinct and μ_1, \dots, μ_n distinct, and let $P_\sigma, P_\tau, P_\delta, P_\gamma$ be the permutation matrices defined by B_2, B_3, B_5 and B_6 (i.e., $B_2 = P_\sigma B_1 P_\sigma^T$, $B_3 = P_\tau B_1 P_\tau^T$, $B_5 = P_\delta B_4 P_\delta^T$ and $B_6 = P_\gamma B_4 P_\gamma^T$). Then, the problem (1) has a solution if and only if*

- (1) $P_\sigma = P_\gamma = P_\tau P_\delta = P_\delta P_\tau$ and
- (2) $\prod_{j \in N_i} \lambda_j = \prod_{j \in N_i} \mu_j$ for each N_i in the commutation partition of the commuting pair of matrices P_σ, P_δ .

Finally, we note that there is a simple corresponding result if B_1, \dots, B_6 are simultaneously diagonalizable by similarity and the diagonal forms meet the conditions of Theorem 19.

References

- [1] H. Flanders, Elementary divisors of AB and BA , Proc. Amer. Math. Soc. 2 (1951) 871–974.
- [2] J. Gelonch, The contragredient equivalence for several matrices: a set of invariants, Linear Algebra Appl. 291 (1999) 37–49.
- [3] J. Gelonch, C.R. Johnson, P. Rubi , An extension of Flanders theorem to several matrices, Linear and Multilinear Algebra 43 (1997) 181–200.
- [4] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.
- [5] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.
- [6] C.R. Johnson, E.A. Schreiner, The relationship between AB and BA , Amer. Math. Monthly 103 (1996) 578–582.
- [7] P. Rubi , J. Gelonch, Doubly multipliable matrices, Rend. Istit. Mat. Univ. Trieste 24 (2) (1992) 103–126.
- [8] R.C. Thompson, On the matrices AB and BA , Linear Algebra Appl. 1 (1968) 43–58.